DOCUMENT RESUME

ED 041 843 | 24 | SP 004 090

AUTHOR	Hall, Gene E.
TITLE	The Instrument for the Analysis of Science Teaching: A System for Measuring Teaching Behavior.
INSTITUTION	Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
PUB DATE	69
CONTRACT	OEC-6-10-108
NOTE	25p.; Report Series No. 19
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.35
DESCRIPTORS	Data Analysis, Elementary School Science, *Interaction Process Analysis, Reliability, *Science Instruction, Video Tape Recordings
IDENTIFIERS	IAST, Instrument for the Analysis of Science Teaching, Science--A Process Approach

ABSTRACT

A system of interaction analysis, the Instrument for the Analysis of Science Teaching (IAST, Part 1), and an accompanying sign system (IAST, Part 2) were developed. Originally used to analyze teaching behavior observed in elementary school classrooms where "Science--A Process Approach" was being taught, IAST was later found to be applicable as a research tool and supervisory aid in other curriculum areas and at other levels. Part 1, a 26-category system based on the 10-category Flanders system, includes 10 student, 14 teacher, and two new categories. Data may be treated for the analysis of various tally records, graphs, ratios, matrices, and tactical findings. Part 2 is a 15-item check list designed to measure some of the characteristics considered to be important in teaching elementary science and "Science--A Process Approach" which could not be determined by part 1 alone. Training in using IAST entails approximately 12 hours of discussion and observation practice. Since the validity of the Scott coefficient PI, the customary statistic for calculating observer reliability, appears to decrease as the number of categories in a system increases, a standardized, nationally available recording of teaching behavior (the Suchman recording of an inquiry session available from Science Research Associates) was used in estimating reliability of observers, demonstrating the possibility of establishing a universal reference for estimates of observer reliability. (JS)

ED041843

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

THE INSTRUMENT FOR THE ANALYSIS OF SCIENCE TEACHING: A SYSTEM FOR MEASURING TEACHING BEHAVIOR

Gene E. Hall

Report Series No. 19

September, 1969

The Research and Development Center for Teacher Education
The University of Texas at Austin

SP004090

The research reported herein was conducted under USOE Contract No. OE 6-10-108 and with the cooperation of The University of Texas Science Education Center.

THE INSTRUMENT FOR THE ANALYSIS OF SCIENCE TEACHING: A SYSTEM FOR MEASURING TEACHING BEHAVIOR

Gene E. Hall

Following the development of interaction analysis by Flanders (1960), many research studies have employed Flanders' ten-category system of interaction analysis. The Flanders' system of interaction analysis was developed for classifying classroom verbal behavior as it relates to classroom climate. However, since interaction analysis provides extensive data not only about the frequency of events but also about the sequence of events, it was readily adaptable to other types of behaviors besides those specifically related to classroom climate. Subsequent interaction analysis category systems have been developed to analyze teaching behavior (Hough, 1967; Parakh, 1968), supervisory behavior (Blumberg, 1968) and various other interactions such as affective (Fuller, 1969).

The proliferation of category systems in research has not been documented in any one source, although a recent publication by Amidon and Hough (1967) does describe some of the more recent category systems. This volume also describes some of the varied research studies that have been carried out with both preservice and inservice teacher training as well as systematic observations in varied types of classroom situations. A newsletter, Classroom Interaction Newsletter, has also been established to disseminate information related to interaction analysis.

The purpose of this paper is not to describe how to use interaction analysis; this has been done by others (e.g. Flanders, 1963; Amidon and Hough, 1967; Cunningham, 1967). The purpose, here, is to describe a

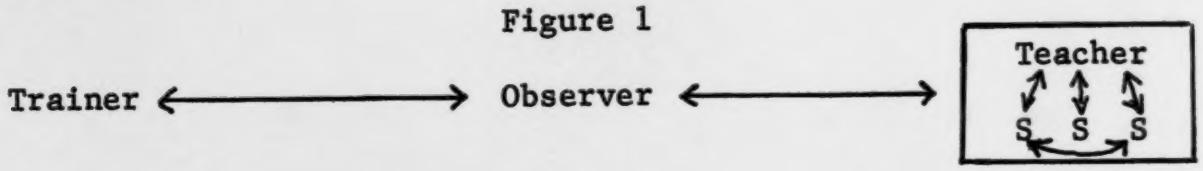
system of interaction analysis, the Instrument for the Analysis of Science Teaching (IAST Part I), and an accompanying sign system (IAST Part II) for measuring teaching behavior. In addition to the description of the IAST, an attempt will be made to describe specific categories and some procedures that others may find helpful in developing systems of interaction analysis for specific research studies and/or supervision work.

Hopefully this paper will also contribute to clarity in the art of communicating findings derived from interaction analysis systems, and also encourage some variety in approaches to category system definition and data analysis.

Before describing the IAST, the various people involved in its use should be introduced. Since the instrument has been developed to observe teaching behavior, the prime people involved in use of the IAST are the teacher and his class of students. A trained observer systematically observes the teacher and students in an instructional situation. Before an observer can use the IAST, training is necessary -- this coming, logically enough, from a trainer. The trainer has not only extensive knowledge of interaction analysis and the types of behaviors to be expected during a classroom observation but also should be sensitive to group dynamics and able to work with a group of observers in establishing ground rules and the "like" mental set that is necessary in order to have inter-observer reliability and validity in the data gathered.

The arrows in the diagram (Figure 1) indicate interactions between the various personnel involved. Hopefully, the observer will witness the instructional situation without causing any, or at least minimal, interference with the session. On the other hand, the trainer and observer(s) should have full interaction with each other. This will contribute to the most reliable and valid information being obtained.

Figure 1



Development of the IAST

The IAST was developed to analyze teaching behaviors observed in elementary school classrooms where Science - A Process Approach (a recently developed elementary school science curriculum) was being taught. However, after the initial development of the IAST, the instrument was found to be applicable both as a research tool and as a supervisory aid in curriculum areas other than science and in other levels of education as well.

During a series of inservice meetings held in the 1965-1966 school year at the Austin, Texas, AAAS Tryout Center, approximately 150 teachers related the approaches and techniques which they had used in conducting daily lessons in Science - A Process Approach. The strategies were then evaluated according to the objectives of the lessons, rationale of the curriculum and the teachers' understanding of the curriculum. From this evaluation, Butts (1966) compiled a list (Key Teacher Ideas) of desired teaching behaviors for teachers employing recently developed elementary school science curricula. Ashley (1967) constructed a sign system (The Classroom Observation Rating Form, CORF) which was designed to record the occurrence of these strategies in the classroom.

The teachers' use of Key Teacher Ideas as well as behaviors listed on the CORF were given prime consideration in the development of the Instrument for the Analysis of Science Teaching (IAST). The IAST has two parts:

I. A 26-category system of interaction analysis;

II. A 15-item checklist (sign system).

The classroom observer, using the IAST, records at least one numerical category from Part I every three seconds based on his observations. If more than one behavior occurs within a three-second interval, then additional categories are recorded. After a prescribed time (15-20 minutes) in which the observer has completed his observations using Part I, he then completes the checklist contained in Part II.

Since its conception, the Instrument for the Analysis of Science Teaching has been utilized as a tool for feedback to preservice teachers after low ratio teaching and in three research studies. College science educators also have been introduced to the IAST as a supervisory tool in their work with inservice teachers. In addition, undergraduate elementary education majors have received training in using the IAST as a form of self feedback through analysis of their own teaching.

IAST Part I

Part I of the IAST is an expanded form of interaction analysis based upon the works of Flanders (1963), Hough (1967) and others. The categories of the IAST Part I account for all of the categories of the original Flanders System and provide more detailed information in certain important areas. The most obvious difference between the IAST Part I and the Flanders System is the expansion of the number of student categories from two to ten categories and the expansion of the number of teacher categories from seven to fourteen. The IAST Part I categories are defined in Figure 2.

The use of "open" and "closed" statements has added a great deal of sensitivity to the IAST Part I over the Flanders System. Guilford, Merrifield and Cox (1961) classified intellectual ability into five major groups: cognition, memory, convergent thinking, divergent thinking and evaluation. Using their operational definitions, a "closed" question is a question requiring a cognitive, memorative or convergent response. In other words, a "closed" question has an accepted right answer (Who discovered America?) while an "open" question encourages divergent or evaluative responses and there is not necessarily a correct, accepted answer (Can you think of a theory to explain this phenomena?). Student statements have also been classified as to being closed (memorative, cognitive or convergent -- "Columbus discovered America") or open (evaluative or divergent -- "It could be due to a presently unknown force field").

Dividing Flanders' Cateogry 3 (accepts or uses ideas of students) into three categories (IAST Part I Categories 2, 3 and 10) has added

FIGURE 2

INSTRUMENT FOR THE ANALYSIS OF SCIENCE TEACHING (IAST)--PART I

TEACHER TALK

0 Student recognition: Verbal or silent recognition for one student to respond.

1 Accepts feelings: Recognizes and identifies with feelings of students (empathetic), non-evaluative encouragement or joking, affective response.

2 Does clarifying of student's ideas: Accepts and uses, or expands ideas of the students. As talk continues change to appropriate category, the restating of the idea of a student, either verbal or written on the board.

3 Causes student to clarify: (non-evaluative) Instead of the teacher clarifying, restating, and/or expanding the ideas of the student, he questions the student's idea and/or asks the student for further explanation.

4 Background or review information (substantive): The teacher gives information from previous lessons or experiences. A restatement of information covered earlier.

5 New information (substantive): Lecturing, facts, calculations, etc., including writing new information on the board.

TEACHER TALK

6 Management information and directions: Giving directions, procedures, telling the student how (what) to do. Requiring an <u>immediate</u> student response or behavior.

TEACHER TALK

7 Asks closed question: A narrow specific, channeled question requiring a specific student response. Application of simple or complex skills to a convergent situation.

8 Asks open question: A broad question, providing space for the student to be original in his response. A "think-type" question.

9 Criticizes or rejects student ideas or behaviors: Self justification and disciplinary statements that may be critical in a defensive manner. A negative value response to student idea. Establishing authority.

TEACHER TALK

10 Gives confirmation (non-evaluative): "yes," "no," "okay," "alright." A short response accepting student's ideas with no value judgement implied. No expansion or clarification of the student's statement.

11 Gives praise: A positive value judgement, words of encouragement.

12 Teacher controlled silence: Of short duration, e.g. a pause after a question. Or a teacher demonstration without talk.

IAST--Part I cont.

STUDENT TALK

13 Student response closed: Student statements that are cognitive, memorative, or convergent in thought.

14 Student response open: Student statements that are divergent, or evaluative in thought.

15 Student affective response: enthusiasm, surprise -- "oh," "gee," etc.

16 Substantive question (closed): A precise, explicit, question about the subject under discussion. A question requiring a cognitive, memorative, or convergent reply.

17 Substantive question (open): Questions that permit divergent or evaluative statements in reply.

18 Procedural question (closed): A question about procedures or directions requiring a specific response about how to do an activity. (cognitive, memorative, convergent)

19 Procedural question (open): An unstructured question about procedures or directions. A question that may be answered with divergent or evaluative statements.

20 Overt activity: Students raising their hands, lab activity, manipulating materials, a group response. This activity must be purposeful. A simultaneous verbal response by several students would be in this category also.

21 Covert activity: Internalized behavior, writing, reading. This activity must be purposeful.

22 Division of student to student interaction: A mark for the separation between two students' interactions.

---

23 Nonfunctional behavior: Behavior without direction or purpose.

5R Teacher reading aloud: Reading a story or information from a book.

13R Student reading aloud: Reading a story or information from a book.

strength to this instrument. There are important differences between the three acts of the teacher restating and/or clarifying a student statement (IAST Category 2), simply accepting the statement with a short yes, no or okay (IAST Category 10) or requesting the student to clarify his statement, "What do you mean by...?" (IAST Category 3).

Category 10 was also employed to provide data about a specific class of teacher responses to student statements. The hypothesis was that perhaps teachers familiar with specific curriculum innovations such as Inquiry Training (Suchman, 1961) would be more apt to use Cateogry 10 type responses.

During observer training concern developed over what to do when a teacher calls on a student. If such an act were classified as a "6" (teacher direction), then the meaning of Category 6 would change. Observers did not feel that it was an easy task to ignore the behavior of calling on students altogether; so Category 0 was established. Note that Cateogry 0 may have very little time dimension, as for example when the teacher names several students in succession to respond to the same question with no additional teacher comment between student statements except a nod or the calling of a name.

Category 6 is only used when the teacher makes a statement requiring <u>immediate</u> student behavior. For example:

T: Go ahead and construct your graphs. (Category 6)

vs. T: Today I am going to ask you to make a graph of an experiment that we will be doing. (Category 5)

Since both teachers and students read aloud in the classroom, the letter R was attached to Category 5 when the teacher reads aloud (5R) and to Category 13 when a student reads aloud (13R).

An hypothesis that in science classes a greater portion of time would be spent in student activity resulted in the identification of two types of student non-verbal behaviors: Overt activity, such as students "bouncing balls" would be classified as 20, while students writing compositions or reading silently would be classified as 21.

Category 22 is a timeless mark indicating that one student talked directly to another student with no teacher behaviors between. Category

22 would only be employed when the student-to-student interaction is at the class level. In other words, it is not used when the students are just gossiping to themselves but only when their statements are a part of, and a contribution to, the ongoing class instruction. This type of behavior appears to be relatively rare; the more common behavior is for student statements to be made to the teacher and responses made by the teacher, rather than one student responding directly to another student with respect to what the first student said.

The Meaning of IAST Data

Once a series of observations have been made using the IAST, tally records of each observation will be available for analysis. A tally record consists of a series of one-minute columns of tallies for the observation period. If the observation period was for 20 minutes, then there would be 20 columns, each column being a tally record for one minute of the observation period. Several types of data analysis are then possible.

The most obvious approach to take would be to examine the tally record minute-by-minute, observing the sequence of tallies and qualitatively observing the frequency of tallies and category repetitions as well as noting any patterns repeating and unusual transitions.

Graphs. A variation on this procedure that may provide some interesting insight into the data would be to plot, by minute, the occurrence or frequency of a given category. For instance, count the number of tallies for Category 9 occuring during each observation minute. During the first observation minute there might be three Category 9's tallied, while in the second observation minute there might be only one. Following this procedure for the total observation will provide data from which a graphical representation of the incidence of Category 9 over the duration of the observation period is pictured. By examining the resulting graph, identification of those points during a lesson at which the teacher and/or the class are losing control over or interest in the lesson is possible. A graph of this type may be accumulated over several observations for the same teacher, or when two different treatment groups are being employed a graph for each treatment group could be constructed and compared.

Another type of graph that can be constructed is a comparative plot of the per cent of the total tallies in specific categories. With a series of bar graphs, visual comparisons are possible of the relative amounts of tallies for several different categories. Two category groupings employed in the past have been: Comparing the percentage of time in the various indirect categories to the percentage of time in direct categories and comparing the percentage of time in teacher categories to the percentage of time in student categories.

Ratios. Another type of data analysis that can be employed is the construction of various ratios. Flanders has suggested the Indirect-Direct ratio, which is better known as the I/D ratio. This ratio provides an indication of the relative amount of time spent in indirect (behaviors which tend to increase student participation) versus direct (behaviors which tend to increase the active control by the teacher) behaviors by the teacher. Various ratios that have been employed when using the IAST are:

The Flanders I/D ratio

$$\frac{\text{Indirect}}{\text{Direct}} : \frac{1+2+3+7+8+11}{4+5+6+9}$$

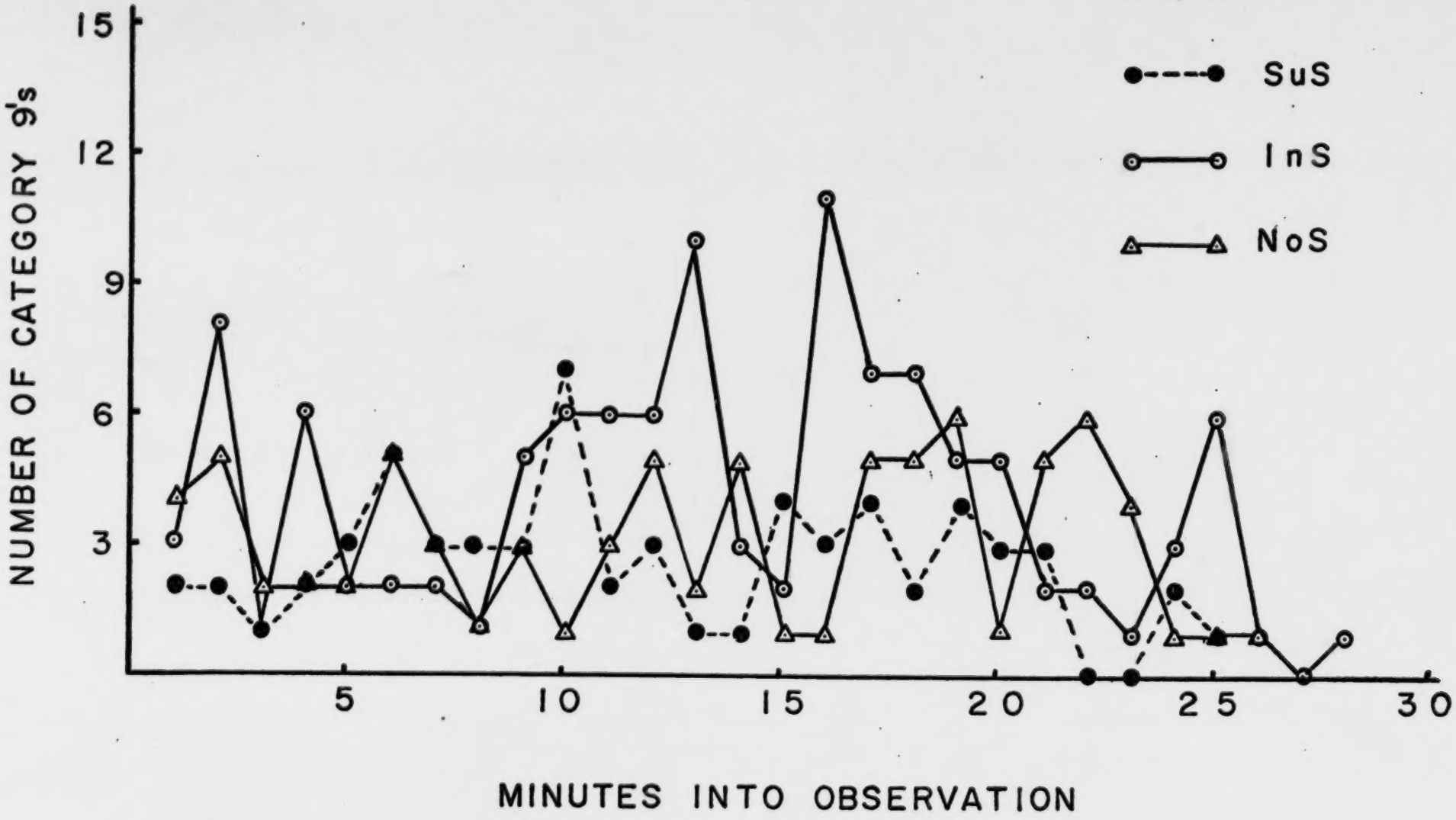
The revised Flanders indirect/direct ratio can also be calculated using the IAST. The revised indirect/direct ratio or i/d ratio provides an indication of the relative emphasis given to direct and indirect motivation and control behaviors during a particular observation(s). This ratio is similar to the I/D ratio except that the substantive behaviors of lecturing and asking questions have been removed.

$$\frac{\text{indirect}}{\text{direct}} : \frac{1+2+11}{6+9}$$

Another ratio that might be of interest is the extended teacher talk-extended student talk ratio. Extended talk involves the continuous use of one behavior for a period of time that is longer than three seconds. This type of behavior has also been referred to as steady state or sustained verbal behavior.

$$\frac{\text{extended teacher talk}}{\text{extended student talk}} : \frac{(1-1)+(2-2)+(3-3)+(4-4)+(5-5)+(6-6)+(7-7)+(8-8)+(9-9)+(11-11)}{(13-13)+(14-14)+(15-15)+(16-16)+(17-17)+(18-18)+(19-19)}$$

FIGURE 3


Group Totals
SuS
InS
NoS
NUMBER OF CATEGORY 9's
15
12
9
6
3
5
10
15
20
25
30
MINUTES INTO OBSERVATION

The total teacher talk, total student talk ratio may also provide interesting insight. This ratio includes all verbal behavior except the teacher calling on a student to respond (Category 0) and student activity (Categories 20 and 21).

$$\frac{\text{Total teacher talk}}{\text{Total student talk}} : \frac{1+2+3+4+5+5R+6+7+8+9+10+11}{13+13R+14+15+16+17+18+19}$$

Another ratio that can be calculated is the number of teacher closed questions divided by the number of teacher open questions.

$$\frac{\text{Teacher closed questions}}{\text{Teacher open questions}} : \frac{\text{Tot 7 - (7-7)}}{\text{Tot 8 - (8-8)}}$$

The per cent of time spent in student closed statements (Category 13) can be divided by the per cent of time in student open statements (Category 14).

$$\frac{\text{Student closed statements}}{\text{Student open statements}} : \frac{13}{14}$$

Another set of ratios that have been of interest are the extended talk/total talk ratios. These ratios are not based on the proportion of talk for the entire observation period that is extended, but that proportion of the total teacher or student talk that is extended. These ratios are:

$$\frac{\text{Extended teacher talk}}{\text{Total teacher talk}} : \frac{(1-1)+(2-2)+(3-3)+(4-4)+(5-5)+(5R-5R)+(6-6)+(7-7)+(8-8)+(9-9)+(11-11)}{\text{Sum of Categories 1-9 and 11}}$$

$$\frac{\text{Extended student talk}}{\text{Total student talk}} : \frac{(13-13)+(13R-13R)+(14-14)+(15-15)+(16-16)+(17-17)+(18-18)+(19-19)}{\text{Sum of Categories 13-19}}$$

Many other ratios can also be calculated. The number of ratios is restricted only by the imagination of the ratio constructor.

Matrix Analysis. Of course, as with any system of interaction analysis, IAST Part I data can be placed into a matrix. However, matrix construction, when 26 categories are being employed -- which means a 26 x 26 matrix -- can be time consuming. A computer program should be constructed to perform this task. The computer program does not have to be very elaborate; although if the programmer becomes enthused, all sorts of helpful additions may be added, such as the total number of tallies for each category and the calculation of ratios.

Matrix analysis follows the traditional method with various zones and regions being related to teacher talk, teacher talk in response to student talk, student talk in response to teacher talk, and so forth. One point of interest, however, is the use of Category 22. This category is employed to mark when one student responds directly to another student without an intervening teacher behavior. By examining the matrix, determination of which student behaviors lead to another student responding (Categories 13 to 19-22) can be made. What the student response is (Categories 22-13 to 19) can also be determined.

Here again, matrix analysis is limited only by the imagination of the reader. As was mentioned earlier, the Category 0 was established to indicate when a teacher calls upon a student to respond, either verbally or by pointing or nodding at the student. From the matrix, since the Category 0 has been recorded, the possibility exists for determining whether the teacher calls upon a student prior to asking a question (Category 0 going to Category 7) or the teacher asks the question and then calls upon a student (Categories 7 or 8 going to Category 0). An hypothesis would be that these two behaviors are different. A disadvantage of having the Category 0 is that the matrix does not have information about what student behavior directly follows a teacher question. By inserting the 0 this information is lost. Therefore, if there is prime interest in knowing whether a student's open or closed statement follows a teacher's open or closed question, then the Category 0 should not be employed or inspection of each individual tally record and the tallying of each special transition as it occurs is required.

Tactical Findings. After examination of the cells for each category of a matrix, a method for identifying the most probable sequence of tallies has been developed (see Figure 5). This procedure is as follows: The single cell of the matrix with the highest number of tallies is identified. This cell represents the behavior (category transition) that occurred most often. For a given category, that cell having the highest number of tallies would be the transition that was most likely to occur. For example, examination of a sample

matrix (see Figure 4) indicated that the 5-5 cell had the highest number of tallies. This means that the single transition that occurred most often was a Category 5 tally followed by another Category 5 tally.

By further examination, the Category 5 row cell with the second highest number of tallies was identified. This cell would be the second most likely transition to occur. The second highest number of tallies was in the Category 5-7 cell. Therefore, the second most probable transition was for a Category 7 to follow a Category 5. Up to this point, then, the sequence of behaviors was for a Category 5 to go to a Category 5, to go to a Category 7. Next, the row cells for Category 7 were examined. By looking for the most probable transition, that Category 7 cell with the highest number of tallies, the most probable category to follow a Category 7 was identified -- in this case, the 7-13 transition. Having compared, by row, the most likely cells for all categories, one might then state the most probable sequence of tallies. This most probable sequence of tallies has been named, by the author, the "primary tactic." The primary tactic which has been identified from studies done in elementary school science classrooms is the following:

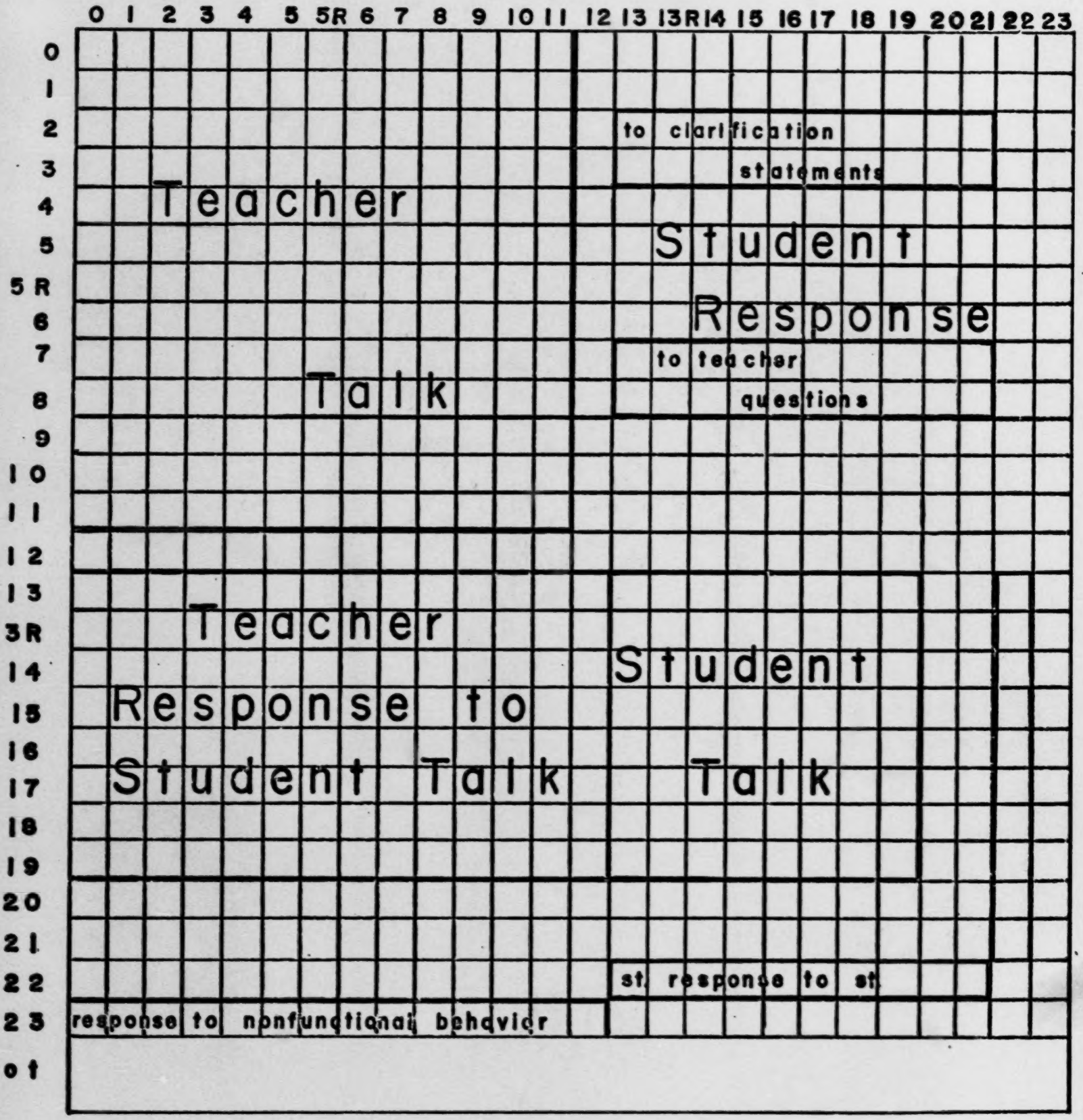
Primary Tactic: a) 5-5-7-13-2-7

This primary tactic consists of extended teacher lecture (Category 5), followed by a short (three seconds or less) closed question (Category 7), followed by a short closed student statement, followed by a short teacher restatement or clarification of the student statement (Category 2), followed by another short teacher closed question (Category 7). In some types of classrooms the primary tactic is modified so that the closed student statement (Category 13) is extended for longer periods than three seconds.

Primary Tactic: b) 5-5-7-13-13-2-7

FIGURE 4

Instrument for the Analysis of Science Teaching (IAST) Part I



Primary Tactic

IAST

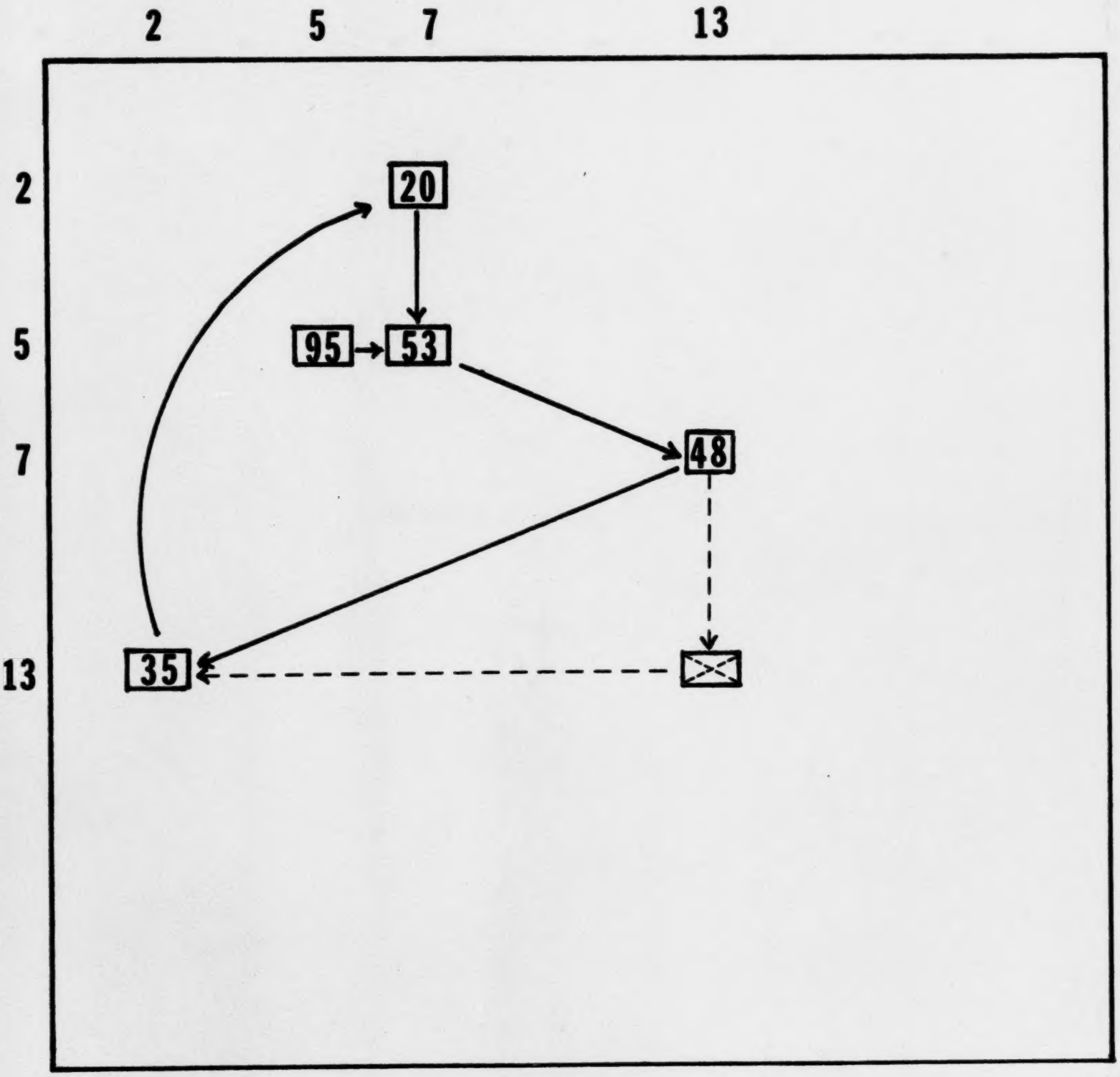


Figure 5

In addition to the primary tactic a secondary tactic has been identified. The secondary tactic being another sequence of tallies for which the total percentage of class time is not as high as for the primary tactic.

Secondary Tactic: 6-6-20-20-6

The secondary tactic consists of an extended set of teacher directions (Category 6 to Category 6), followed by extended student overt activity (Category 20 to Category 20), followed by teacher directions (Category 6).

By adding the total number of tallies that accounted for each of the transition cells included in an identified tactic and dividing by the total number of tallies for the observation, a calculation of the percentage of observation time that is accounted for by a tactic may be made. In second-grade classes employing a recently developed science curriculum, an average of 24.2 per cent of the class time was spent in primary tactic transitions, while an average of 15.0 per cent of the class time was spent in secondary tactic transitions.

IAST Part II

Part II of the IAST (Figure 6) is a sign system consisting of 15 items to which the observer responds on a seven-point semantic space. These items are designed to measure some of the characteristics considered to be important in teaching elementary science as well as characteristics important to teaching Science - A Process Approach (Butts, 1966; Ashley, 1967) which could not be determined by using Part I alone.

At the end of an observation period the observer completes Part II, including the final item which refers to demographic data about the instruction period. An item is not checked unless there has been a definite applicable behavior witnessed during the observation period. A response in the neutral zone of the semantic space is different from the absence of the characteristic all together.

FIGURE 6

OBSERVER'S CHECKLIST IAST PART II

Teacher:_____ Date:_____(p.m., a.m.)(M T W T F)

School:_____ Observer:_____

Item 1-15 are to be completed at the conclusion of each observation period. For each item, indicate along the continuum the point that best represents the observation period just completed.

Room

1. Room attractively arranged with evidences of provoking materials for learning.

many _____:___:___:___:___:___:_____ none

2. Display of materials.

teacher produced only _____:___:___:___:___:___:_____ student produced only

3. Proportion of display materials that are science oriented.

all _____:___:___:___:___:___:_____ none

Teacher

4. Presentation of lesson.

good _____:___:___:___:___:___:_____ poor

5. Illustrations of an idea are selected progressively from simple to less obvious.

always _____:___:___:___:___:___:_____ never

6. The teacher has objectives of lesson clearly in mind.

not at all _____:___:___:___:___:___:_____ completely

7. Experience before vocabulary

never _____:___:___:___:___:___:_____ always

8. Pupil activities rather than lecture, or reading.

never _____:___:___:___:___:___:_____ always

9. Teacher works with students at their own rate of speed.

always _____:_____:_____:_____:_____:_____ never

10. Teacher evaluates student performance on basis of what students are able to do.

always _____:_____:_____:_____:_____:_____ never

11. Teacher shows willingness to admit lack of knowledge or instances of error.

never _____:_____:_____:_____:_____:_____ always

12. Teacher shows flexibility in event of problem.

always _____:_____:_____:_____:_____:_____ never

Student

13. Students are enthusiastic toward lesson.

never _____:_____:_____:_____:_____:_____ always

14. Students listen when other students verbally or physically participate.

never _____:_____:_____:_____:_____:_____ always

15. Number of students being involved.

all of them _____:_____:_____:_____:_____:_____ none

Comments

Lesson taught, type of activity, length of observation, use of A.V., seating arrangement, etc.

For data analysis, the semantic space for each item from the IAST Part II can be divided into six parts. The parts are then successively numbered from 1 through 6 from left to right. For each item the space marked by the observer is assigned the numerical value of that part. Then the average value for that item is determined by taking the arithmetic average over the number of observations of a teacher or teachers. Means can then be compared, although deviation from means might provide more information.

The IAST Part II does not appear to be sensitive enough to be employed in research studies. Observers do not seem to be able to rate classrooms on a relative basis with any degree of discrimination. However, possibly with work, its inadequacies can be repaired. The IAST, Part II does have implication for supervisory work, though.

Training of IAST Observers

Training in using the IAST will, of course, vary depending upon whether the user will be employed in a supervisory capacity or a research study. Basic training in using the IAST would normally entail approximately 12 hours of discussion and practice using audiotapes, videotapes and, if possible, live classroom observations. During this training period, the observers would of course have to memorize the category numerals and definitions and become expert in recording a tally approximately every three seconds. If the observers are to be employed in a research study, training should continue beyond the initial 12 hours with additional practice sessions using audiotapes, videotapes and live classroom visits. During the training session observers should discuss and come to mutual agreement and understanding on how to classify each situation. All observers should be present at each training session. If the observers are not always present at the meetings they may not have "like mental sets" with regard to the classification of observed behaviors, thereby reducing reliability and validity.

During the training period, a series of ground rules should be established in order to further clarify and refine the observer "mental set." The use of ground rules will also help the validity and reliability of the use of the instrument, since all observers will be classifying each behavior more nearly the same. Figure 7 is a list of ground rules that have been found useful in using the IAST.

Observer Reliabilities

Observer reliabilities should be calculated from time to time during the training period since this provides feedback to the trainer as well as to the observers, indicating their success in classifying behaviors. Observer reliability should also be checked from time to time following the initial observer training period since observer agreement may regress with time. For this reason, after the training period, weekly or biweekly meetings of observers should be held throughout the observation period to clarify any observer problems and answer any questions.

The Scott coefficient PI (Flanders, 1963) is the customary statistic for calculating observer reliability. The coefficient is based upon the probability of the proportion of tallies in each category being the same for two observers. Unfortunately, the Scott coefficient only provides a partial reflection of the true observer reliability. Several factors effect the value and validity of this coefficient.

Since the calculation of PI is based on differences in the proportion of tallies in each category, if two observers have the same number of tallies for a category but the total number of tallies over all categories differ, then there will be a difference in the proportions for a category resulting in a lower coefficient. With a large number of categories (the IAST Part I has 26 categories), relatively small proportion differences for each category can accumulate to a sufficient sum over 26 categories to significantly reduce the coefficient.

FIGURE 7

GROUND RULES FOR PART I OF THE IAST

1. Classification of teacher statements should be in terms of what the pupils perceive the intent to be -- not what the observer infers the teacher to be saying.
2. Pupil statements should be classified in terms of how the teacher perceives the intent.
3. Teacher talk overrules student talk.
4. If more than one student is talking, record the student that the teacher responds to.
5. Be careful of verbal habits without meaning when using Categories 2, 10 and 11.
6. Category 4 is used only when information from a previous lesson is stated.
7. Category 6 is used when a statement requires immediate student behavior.
8. Use Categories 7 and 8 only when questions are stated as questions.
9. Category 10 is used only in response to student statements or behaviors and that is all the responses made to that student's statements.
10. Categories 20 and 21 and the open and closed statements and questions will be classified differently at different grade levels; these statements and activities are related to the maturity of the individual.
11. If you are uncertain about a teacher statement, classify it as a 6.
12. If you are uncertain about a student statement, classify it as a 13.

Another factor affecting the coefficient is the degree of difficulty in categorizing the behaviors being observed. Relatively high coefficients are easily obtained if few categories are used and there are few transitions in the behaviors being classified (e.g. extended teacher lecture). However, with short statements and many transitions, two observers will be less likely to have identical records.

An additional factor affecting observer reliability is the problem of quick changes in observed behavior -- two or possibly even three tallies for a given three-second interval. Such rapid transitions result in variations in the three-second interval each observer is using, as well as increasing the possibility of an observer missing one or more transitions.

As the number of categories in a system of interaction analysis increases, the validity of the Scott coefficient appears to decrease due to the above mentioned factors. With the IAST Part I, PI coefficients of .70 to above .90 have been obtained depending upon the difficulty of the behaviors being observed and the competence of the observers.

In order to improve communication of observer reliability information, a standardized, nationally available recording of teaching behavior might be adopted. One such recording that has been used in estimating reliabilities of observers using the IAST is the Suchman recording of an inquiry session available from Science Research Associates (Suchman, 1966b). The use of this commercial recording makes possible comparisons of observer reliabilities in one study with other observers in future studies as well as establishing a basis for comparison by other researchers. All of the IAST categories are not utilized in this recording; however, sample behaviors of 15 of the 26 categories are demonstrated. If this recording is unfamiliar to observers -- is only played for reliability check sessions and there is no discussion of interpretations of this specific recording -- then the recording is not apt to be learned by the observers. In other words, reserve this recording for actual reliability checks. Scott coefficients for three observers trained in using the IAST for a recent research study ranged from .70

to .86 with an average of .77 for four 20-minute reliability checks over a six-week period. If this, or a similar recording were used by other researchers, then a basis for comparing interaction analysis system reliabilities would be available.

Conclusions

The development of another system of interaction analysis in and of itself is probably not of profound importance. However, the identification of an approach to measuring reliability with the use of a commercially available audiorecording, the use of a sign system along with a category system, the naming of one or two new categories and the description of some different types of data analysis hopefully have justified this description of the Instrument for the Analysis of Science Teaching. Obviously Parts I and II of the IAST are not final, ultimate answers to the construction of instruments for the analysis of teaching behaviors. However, the results of the use of the IAST, along with other recently developed interaction analysis systems, may lead to the development of more sensitive instruments. At least the IAST has demonstrated the feasibility of using a 26-category system of interaction analysis; and, in the use of a commercial recording, the possibility of establishing a universal reference for estimates of observer reliability has been demonstrated.

BIBLIOGRAPHY

Amidon, E.J., and Hough, J.B. Interaction Analysis: Theory, Research and Application. Reading, Massachusetts: Addison-Wesley, 1967.

Ashley, J.P. A Study of the Impact of an Inservice Education Program on Teacher Education. (Doctoral dissertation, The University of Texas), Austin, 1967.

Blumberg, A. A System for Analyzing Supervisor-Teacher Interaction. Unpublished paper, Temple University, 1968.

Butts, D.P. The Teachers' Ideas. Science Inservice Project Manual for Teachers. Austin: The University of Texas, 1966.

Classroom Interaction Newsletter. Research for Better Schools, Inc. Philadelphia.

Cunningham, J.D. Interaction Analysis: A Useful Technique for Research and Science Supervision. Science Education, 51: 27-32, Feb. 1967.

Flanders, N.A. Interaction Analysis in the Classroom: A Manual for Observers. Ann Arbor, Michigan: School of Education, University of Michigan, 1963.

Fuller, F.F. FAIR System Manual, Fuller Affective Interaction Records. Research Methodology Monograph No. 6. The Research and Development Center for Teacher Education, The University of Texas at Austin, June 1969, 103 pages.

Guilford, J.P., Merrifield, P.R., and Cox, A.B. Creative Thinking in Children at the Junior High Level. (U.S. Office of Education Cooperative Research Project No. 737) Los Angeles: University of Southern California, 1961.

Hough, J.B. An observational system for the analysis of classroom instruction. In E.J. Amidon and J.B. Hough (Eds.), Interaction Analysis: Theory, Research and Application. Reading, Massachusetts: Addison-Wesley, 1967. Pp. 307-314.

Parakh, J.S. Teacher-pupil Interaction in Biology Classes. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Chicago, February 1968.

Suchman, J.R. Developing Inquiry. Chicago: Science Research Associates, 1966.

RECORDING

Suchman, J.R. An Inquiry Session. Chicago: Science Research Associates, 1966b.